

## DOCUMENT RESUME

ED 411 834

IR 056 661

AUTHOR Hanani, Uri; Frank, Ariel  
TITLE Intelligent Information Harvesting Architecture: An Application to a High School Environment.  
PUB DATE 1996-00-00  
NOTE 12p.; In: Online Information 96. Proceedings of the International Online Information Meeting (20th, Olympia 2, London, England, United Kingdom, December 3-5, 1996); see IR 056 631.  
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)  
EDRS PRICE MF01/PC01 Plus Postage.  
DESCRIPTORS Access to Information; Computer System Design; Computer Uses in Education; \*Electronic Libraries; Elementary Secondary Education; Foreign Countries; High Schools; \*Information Retrieval; \*Internet; \*School Libraries; Technological Advancement  
IDENTIFIERS Digital Technology; \*Israel; Search Engines

## ABSTRACT

In the educational arena, information is conventionally scattered throughout many projects and documents and on many systems. This distribution of data inhibits students and faculty members from searching and accessing information conveniently and efficiently. The research project described in this paper aims to consolidate the disparate data into one information repository. Known as the KATSIR (K12 Advanced Touring System based on Information Retrieval) system, the project is developing and implementing a comprehensive architecture for intelligent information retrieval in open systems. The novelty of this approach is the combination of a new research paradigm in information retrieval, called information harvesting, with a K12-friendly interface. This paradigm enables both teachers and students to gain practical experience in harvesting information both locally and throughout Internet sites in a K12 environment. As part of this research, an innovative information retrieval project was developed. The program is targeted mainly at the establishment and implementation of a comprehensive Educational Digital Library. This new virtual school library was implemented in the Gilo Comprehensive High School on a local area network that contains more than 150 personal computers with a CD-ROM based system, a high-speed line interface to the Internet, and advanced information science tools. This paper presents the KATSIR system with its various components and capabilities on the Internet as a powerful search and harvesting engine, and its promising contributions to the educational environment. (Contains 12 references.) (Author)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

# **Intelligent Information Harvesting Architecture: An Application to a High School Environment**

**By:**

**Uri Hanani**

**Ariel Frank**

"PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY

B.P. Jeapes

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

**BEST COPY AVAILABLE**



# Intelligent information harvesting architecture: an application to a high school environment

Full Text Provided by ERIC

B. P. Jeapes

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION

CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

**Uri Hanani and Ariel Frank**

*Bar-Ilan University, Israel*

**Abstract:** In the educational arena, information is conventionally scattered throughout many projects and documents and on many systems. This distribution of data inhibits students and faculty members from searching and accessing information conveniently and efficiently. This research project aims to consolidate the disparate data into one information repository. Known as the KATSIR system, the project is developing and implementing a comprehensive architecture for intelligent information retrieval in open systems. The novelty of this approach is the combination of a new research paradigm in information retrieval, called information harvesting, with a K12-friendly interface. This paradigm enables both teachers and students to gain practical experience in harvesting information both locally and throughout Internet sites in a K12-environment. As part of this research, an innovative information retrieval project was developed. The programme is targeted mainly at the establishment and implementation of a comprehensive Educational Digital Library. This new virtual school library was implemented in the Gilo Comprehensive High School on a local area network that contains more than 150 personal computers with a CD-ROM based system, a high-speed line interface to the Internet, and advanced information science tools. This paper presents the KASTIR system with its various components and capabilities on the Internet as a powerful search and harvesting engine, and its promising contributions to the educational environment.

**Keywords:** Internet, information retrieval, open systems, intelligent information harvesting, computers in high school, multilingual (Hebrew) support

## 1. Introduction

Due to the huge volume of information gathered on the Internet and its heterogeneity, search engines were developed to enable users to find their way in the World Wide Web. More specifically, in the educational arena, the information is often scattered throughout many projects and documents, and on many systems. As a result of the recent trend of using more computers in school, faculty members prepare their background materials on the computers, and pupils surf the Web as part of their research and doing homework.

Moreover, the transition from frontal teaching to a teaching environment based on information retrieval and telecomputing poses a new challenge. One can note a 'Tower of Babylon' of information resources and educational work that is scattered throughout separate projects and documents. In other words, this distribution of data inhibits students and faculty members from searching and accessing information conveniently and efficiently.

This problem exists primarily in a mature environment of K12 institutions that are pioneers in the introduction of computers to the educational arena. One such pioneer is the Gilo Comprehensive High School at Jerusalem, Israel.

In the early 1990s, the Israeli Ministry of Education and Culture appointed a committee for telecomputing schools all over the country. In 1992, the committee submitted a proposal for integrating computers into the education process in Israel. Since then, much effort has been made to increase the use of computers at schools as a part of a project known as 'Tomorrow 98'. As part of this strategic plan, ten schools were chosen to represent various models of implementing telecomputing environments throughout this project.

The Jerusalem Gilo Comprehensive high school, as one of these schools, has made an effort to construct a model of a virtual library school based on information retrieval, information science paradigms and telecomputing environments. This advanced new virtual library is implemented on a local area network that includes more than 150 PCs, with a CD-ROM server-based system, and a high-speed line interface (frame-relay) to the Internet. The networked system also supports hypermedia authoring tools, a sophisticated video studio and advanced information science tools such as SQL Engine, HTML and harvesting tools. This existing technical know-how and

capabilities enable the teachers and pupils to take advantage of sophisticated search strategies for locating relevant information from the local networks and on the Internet, and to develop advanced Web home pages, Web hypermedia presentations and useful courseware. The full view of Gilo Comprehensive is partially reflected at the Gilo home page. There, some of multimedia projects can be viewed and retrieved. The URL is <http://www.gilo.jlm.k12.il>.

## 2. Digital libraries and the virtual school

Apart from the pedagogic innovations introduced through advanced information technologies in the Gilo school model, its management tackled the above-mentioned problem of abundance of dispersed documents and media. It was evident that a new coherent understanding was needed to encounter the dilemma involved in the collection and management of digital materials. Here, as a combination of academic know-how and real life experience, the technology recently termed 'virtual digital library' was applied (Ref 1).

One has to comprehend that the concept of a digital library (DL) in a virtual school expresses the revolution in education that has emerged from information technologies and telecomputing. This can lead one day to an environment that is free from the constraints of time and space. Teachers and students can have interactions without attending specific classes according to a rigid schedule. The connection between teachers and students is achieved by means of telecomputing systems based on communication networks. Moreover, most information resources are not limited to the conventional school or library. Computerised databases and communication networks make access to information possible for anyone from anywhere at all times.

This future (or actually the present environment in our case) invokes some new possibilities (as well as questions) such as:

- the role of the teacher as a mediator instead of being a classical owner of knowledge (Ref 2);
- augmentation of the traditional librarian with machine-oriented 'intelligent agents' (Ref 3);
- new methods of human-computer interactions rich with assistance and guidance (Ref 4, 5).

Based on this technologically advanced environment, many DL projects have been initiated at the Gilo High School, to the benefit of both students and teachers. Using the DL, students are able to consolidate several related projects into one information repository. The system also gives students the opportunity to access Hebrew documents, regardless of format, through a Hebrew interface. Ordinarily, this poses a problem due to different computer operating systems, Hebrew formats and general interoperability problems. The infrastructure constructed can also be enhanced to support a multilingual environment (besides English and Hebrew). Another important property of the DL is its 'K12-friendly' interface that is designed to make high school students comfortable and enthusiastic about conducting searches and viewing their results.

To support all these requirements of the educational environment, an innovative project was initiated to develop and implement a comprehensive architecture for intelligent information retrieval in open systems. The novelty of this approach is the combination of a new research paradigm in information retrieval called information harvesting with a K12-friendly interface. This paradigm enables both teachers and students to gain practical experience in harvesting information both locally (intranet) and throughout Internet sites within the K12 environment. The Gilo project is targeted mainly at the establishment and implementation of a comprehensive educational digital library, or in other words a virtual school library. The educational information retrieval system operating at the school is based on the Harvest system originally developed at Colorado University (Ref 6) and is titled KATSIR — 'K12 Advanced Touring System based on Information Retrieval' ('katsir' in Hebrew means 'harvesting').

## 3. KATSIR description

We give here a concise and brief (and slightly technical) description of the KATSIR project that is the major theme of this paper. We start by a short survey of the Harvest system that served as one of the starting points of our research.

### 3.1. The Harvest approach — information gathering and access

The Harvest project was launched at the start of the 1990s at Colorado University (Ref 7). It operates as a server on the Web and is targeted to achieve three main goals:

- (1) an infrastructure architecture that can collect (Harvest) distributed indexed information from the Internet, in an efficient manner, with minimal overload on the networks and communications channels;
- (2) detailed customisation of different sorts of indexes over a wide spectrum of varied databases, heterogeneous schema, information resources and URLs;
- (3) support for local caching and information replication, to provide fast response time and the sharing of computing resources.

The Harvest system consists of six components:

- (1) **Gatherer:** deals with the collection of indexed data from the Internet using local providers in an incre-

mental way, and in sending it from the local cache in a compressed stream. It is based on the Essence subsystem that specialises in different data types formats and their retrieval. Essence deals with summarisation, after scanning the original document, by extracting important information from the original document. In this syntactic phase, Essence can be guided to 'understand' and filter significant knowledge and catalogue it along with the indexing process. In HTML format, the different parts are annotated by such special tags as Header, Title or Bold. The Gatherer can be directed to carry on its branching search from URL to URL, or to stop after several link traversals.

- (2) **Broker:** supports a user query interface on the Harvest host over the collected indexed data. Different brokers can operate in parallel, retrieving information from their own host or from all brokers in the harvest network. The brokers update their indexes in an incremental manner whenever a query is invoked or data is updated. A special Broker HSR (Harvest Server Registry) is responsible for the information about all brokers, Replicators and Caches in the entire system.
- (3) **Index/Search Subsystem:** a general interface to Internet search engines (such as WAIS, Glimpse, Nebula). It supports Boolean search and incremental updates. Glimpse is equipped with an efficient indexing system and interactive queries are used as the standard user query interface.
- (4) **Replicator:** replicates data in a weakly-consistent manner over distributed file systems in the network. The replicated data is kept in the network with mirror copies using suitable protocols.
- (5) **Object cache:** its main aim is to manage the system caching memory, to optimise searching for files and data over the network, and save redundant accesses to retrieve data.
- (6) **Object system:** handles complex types of objects that are kept in the network. It allows the storage and the retrieval of objects from local and remote hosts.

Summing up, the Harvest system is constructed as an open and scaleable architecture using Resource Discovery methods. It is customisable, flexible and adaptive to many applications. More technical information and a list of Harvest sites and uses can be found at Ref 7 and at the URL <http://harvest.cs.colorado.edu/>.

### 3.2. KATSIR objectives and phases

The objectives of the KATSIR project outlined in the paper are as follows:

- the development of a representative model of referential and abstracts materials (Educational Repository) in the Gilo digital library, with specific emphasis on K12 educational documents;
- using the HARVEST Internet search software developed at the University of Colorado to develop and implement an open architecture for intelligent information retrieval, using a friendly K12 human-computer interface, while applying methodologies of 'information brokers', 'intelligent agents' and harvesting tools;
- evaluating the prototype developed by applying it to an educational repository including the drawing of conclusions, from students and faculty use and feedback, that can be relevant for the entire educational community.

To carry out this pioneering project, several phases were outlined:

- investigating the Harvest system and its tools;
- building the prototype educational information data repository as a Digital Library;
- developing a K12-friendly user interface;
- testing and evaluating the prototype with student and faculty feedback.

Before describing these phases and their specific results, we note that the KATSIR architecture framework main contribution is to enable high school students to feel a sense of accomplishment when accessing information and computer technologies. Students are encouraged to gain skills in investigating information systems and in using practical applications of various computer technologies.

## 4. KATSIR components

KATSIR is composed of four main processes, as follows:

- (1) the collection (gathering) of initial K12 Internet relevant information and URLs;
- (2) information processing and summarisation;
- (3) the presentation of processed information and its retrieval;
- (4) the analysis and follow-up of system usage.

### 4.1. Initial K12 collection

The gathering process is activated by applying a Harvest Gatherer to an Internet K12 relevant URL. The initial URLs are provided manually to the Gatherer by information scientists or by system users. Hereafter, the Gatherer follows the HTML links from these URLs and collects the relevant documents using recursive branching.

This automatic branching process can be controlled by the system, either by instructing it to stop after a

specific number of branches or by a filtering algorithm (this filtering process is still in a development phase and evaluation).

In contrast to the Harvest original Gatherer, our improved Gatherer is capable of collecting documents that reside on the local user network (high school LAN). Here, using a special KATSIR user interface, the student or teacher can update the database with relevant material such as multimedia presentations or any other relevant documents. This data entry option allows the inclusion of the document description and storage location in the LAN. Thus, the DL is enriched with many interesting materials from the local school environment.

## 4.2. Information processing and summarisation

Here, four sub-processes are available:

- (1) Summarisation;
- (2) Hebrew support;
- (3) Optimisation and cataloguing;
- (4) Indexing.

- **Summarisation.** Here the Harvest Essence tool is applied. The original collected document is scanned and several important details (such as Document Title, URL, short description, Key Words and a few lines from the document initial content) are registered by use of an abstracting algorithm. This abstraction process is syntactically oriented. However, as mentioned before Essence can be guided to give a special treatment to HTML tags. It should be noted that in order to save disk storage, the original collected document is not kept in the database, but it can be found using its original URL in the Internet.
- **Hebrew support.** Native English speakers are usually not aware of the problems caused by the Semite languages (Hebrew, Arabic) in the information retrieval arena. Due to writing and reading from right to left (in contrast to Western languages), special treatment must be given to documents that are written in Hebrew. KATSIR supports Hebrew and is capable of converting the document summary, abstracted by the previous phase (summarisation), to a special format that is compatible with the user interface (which operates either in Windows mode or other GUI systems). This necessitates a set of special Hebrew algorithms that deal with the Hebrew orientation indexing and retrieval parts in KATSIR.
- **Optimisation and cataloguing.** Due to the KATSIR educational environment, we developed further the properties related to filtering, optimisation and cataloguing of the documents. After the automatic summarisation done by Essence, the information scientists can:
  - delete or add keywords to the document;
  - correct any field or description annotated by the system;
  - decide whenever the document is relevant or not;
  - attach any document to the system browsing outline tree that guides the users in their queries.
- **Indexing.** The original document summaries are processed to produce indexes keys that serve the search engine. KATSIR uses the Harvest standard search engine, namely Glimpse.

## 4.3. The presentation of processed information and its retrieval

KATSIR allows two main interaction styles (or access methods): information *retrieval* and *touring* (i.e. browsing). The retrieval of information follows Harvest standard text retrieval mechanisms that were tailored to the KATSIR environment. The second method is one of KATSIR's main innovations. Here, the user can take a tour along a topics tree that represents the various educational entry points for both students and teachers. The system automatically generates an HTML page which displays the topics tree that contains links to the relevant materials.

In order to simplify the system usability for the typical K12 user, the user is advised to make use of Glimpse's full (and complicated) search options by special branching. In parallel, KATSIR also advises the user to conduct a simple structured search by a 'frame text'.

## 4.4. The analysis and follow-up of system use

All activities in KATSIR are logged and monitored, including users' touring steps. This will allow analysis of system usability and performance, and may suggest further research.

# 5. KATSIR applications

## 5.1. System integration and evaluation

A main objective of the KATSIR project direction is its K12 orientation. We consider its usability for a typical K12 user, either faculty or student, to be one of the key issues. So, system integration and evaluation were carefully

planned to evolve as follows:

- (1) system analysis and users' requirement definition;
- (2) prototyping and pilot;
- (3) system evaluation and test runs;
- (4) further development;
- (5) further evaluation and refinement.

System evaluation and tests were made by special teams of students and faculty called 'leading groups'. These groups were consolidated during a two-year process of intensive work, as part of the Gilo representative model, and they also serve other objectives in the school modeling environment. For example, the Faculty leading group chose the information domains that should be included in the topics tree, and they gave the initial feedback to the documents and URLs that were gathered by the research group. In addition, the leading students participate intensively in the evaluation of the effectiveness of the KATSIR query interface. Without these contributions, KATSIR would still be just an academic tool.

## 5.2. Applications

Taking advantage of the Gilo school daily involvement in all the above-mentioned phases, the students and faculty members are part and parcel of the KATSIR development. The many users involved and the educational projects that were integrated into the DL were just results of this approach.

To illustrate what was achieved in this pedagogic domain, we list here some of the educational projects that are part of these activities as DL components (see Appendix):

### ***Students' projects***

- specials events in Israeli political life and atmosphere;
- information banks such as: stamps, NASA activities, participation in special education activities and work groups;
- designing tutorials for other students in the domains of: HTML, 3Dstudio, multimedia tools, etc.;
- electronic newspaper, updated bi-weekly, related to school life as well as politics, current affairs, etc. This project includes also a reference to two scientific electronic newspapers available in Hebrew.

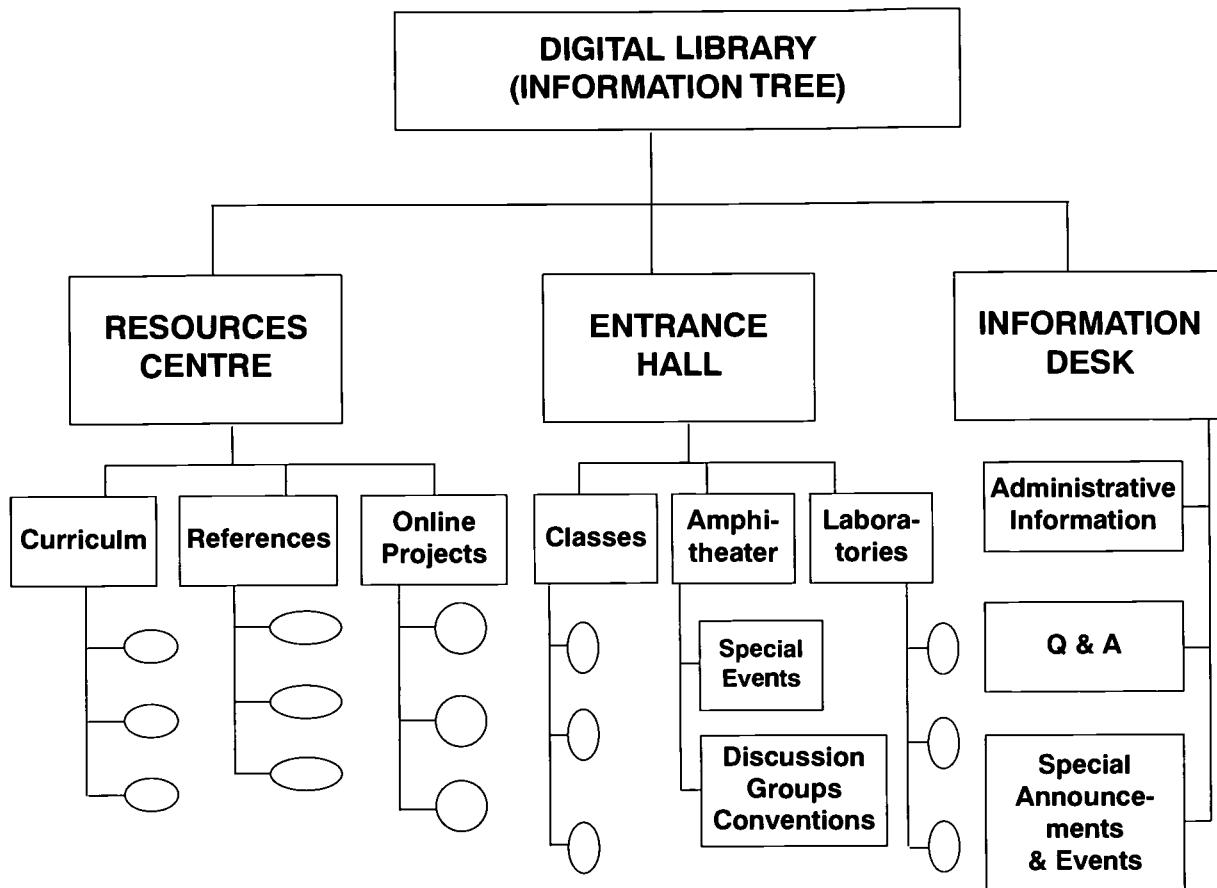
### ***Information for the faculty***

- International projects of K12 teachers and discussion groups;
- multimedia presentations done by colleagues and even students that served as reference materials for classes, in many domains such as:
  - physics;
  - humanities;
  - history sciences;
  - geography;
  - holocaust studies;
  - mathematics, geometry and trigonometry;
  - biology.

### ***Internet section***

- HTML tutorials;
- Unix tools and tutorials;
- repository of graphical aids to prepare home pages and converters;
- Q&A material and FAQs.

In Figure 1, we try to illustrate all digital library materials in their natural educational environment. The objective is to represent all school activities as one complete information tree, representing a digital library. This metaphor is user friendly and simplifies the DL touring, while making the browsing and retrieval easy for faculty and students. This ongoing sub-project is evolving as a result of the above-mentioned process of evaluation and refinement. We have a future plan to convert it to three-dimensional space, using VRML tools.



**Figure 1:** Digital library as a metaphor of K12 environment.

### 5.3. Local library — intranet

In the analysis study phase we have identified a typical problem that characterises the K12 sector. Not every school can afford an Internet connection, as is possible in the Gilo case. So, an intranet structure (i.e. a local Internet TCP/IP network without immediate access to the Internet) can be installed on the LAN. One essential Intranet tool is a search engine that supports a local DL. Therefore, an SQL engine was tailored to work with KATSIR as an additional tool that handles local digital materials and documents. This tool contains a simple SQL search engine and a topics tree, with a user guided interface, as illustrated (partially) in Figure 2. This local DL can serve as an intermediate repository to the KATSIR DL, and materials can be collected and exported from this repository as from any other Internet repository.

NetScape [Galo's Internet Project]

File Edit View Go Bookmarks Options Directory Window Help

Back | Forward | Home | Reload | Images | Open | Prev | Find | Stop |

Location: <http://www.galo.jm.k12.il/educate/index.ctl>

What's New? What's Cool? Destinations Net Search People Software

**Search Database**

**Step One:**  
Please pick options from the categories below. If you don't want to limit yourself in your search choose the "any" option

<b>Humanist</b>	<b>Science</b>	<b>Language Studies</b>
<input type="radio"/> Literature	<input type="radio"/> Physics	<input type="radio"/> English
<input type="radio"/> Grammar	<input type="radio"/> Chemistry	<input type="radio"/> Arabic
<input type="radio"/> Bible	<input type="radio"/> Biology	<input type="radio"/> Hebrew
<input type="radio"/> History	<input checked="" type="radio"/> Mathematics	<input type="radio"/> ALL
<input type="radio"/> Government	<input type="radio"/> Electronics	
<input type="radio"/> Geography	<input type="radio"/> Computer Science	

**Step Two**  
Choose the type of work

All Works

**Step Three**

Figure 2: Search engine SQL based digital library repository.

## 6. Further Developments in KATSIR

KATSIR was initiated as an open infrastructure architecture for intelligent information retrieval and touring. Therefore, it is natural that some innovative developments of KATSIR are under work at the present. We list here some of these developments:

- *Intelligent human-computer interface* that includes a thesaurus and consulting facilities, based on the expertext ideas presented at Hanani (Ref 5), Rada (Ref 8) and Chen et al. (Ref 9);
- *Developing the Harvest collection process* (the Gatherer) by applying new filtering algorithms based on ideas developed by Shapira et al. (Ref 10) while researching user filtering;
- *Developing a three-dimensional human-computer interface* based on VRML and Java applets, using the metaphor of school environment and school layout;
- *More advanced ranking* of search results presented in a friendly way to the K12 users.
- *Applying the KATSIR infrastructure* to regional or even country-based environments (such as urban Jerusalem or Israel), by interconnecting several brokers with a single KATSIR management system. This will allow the application of the caching mechanism of Harvest, with optimisation of search accesses to the Internet, and really arriving at a community virtual school and virtual digital library.
- *Applying a more advanced usability evaluation* other than the user-task model (see Ref 11), and studying advanced productivity tools in the digital library arena (see Ref 12).

## 7. Conclusions

In this paper we presented the innovative approach of the KATSIR system, based on the Harvest architecture, as an Internet collection system and digital library applied to the educational environment. One major achievement is the cooperation between the academic research team and K12 students and faculty members that allowed the realisation of an educational digital library operating in the field. Our KATSIR environment is an open architecture

that can form a generalised infrastructure to be used successfully by any K12 schools and organisations.

The Internet is going to be the main issue in information science, both as an academic theme and as an application technology. The information gathered at the many (million or more) URLs, home pages and sites will not be available to the user community unless we pay attention to user needs and develop a more adaptable paradigm that will replace the old notion of search engines. Here, the KATSIR project serves as a model as to what that can be done successfully without a large budget and with real life applications. Thus, we have enough background and experience to believe that this research is an important contribution to the information retrieval community of new and beneficial tools and concepts, that can be put to use now and in the near future.

## Acknowledgements

This research was funded by the Israeli Ministry of Science and Arts, in the framework of experimentation and demonstration in the domain of information superhighways, digital library and data repositories. The authors are also indebted to I. Miller, KATSIR Chief Knowledge Engineer, and the rest of the KATSIR research group of Bar-Ilan University, Ramat-Gan and the Gilo High School, Jerusalem, Israel.

Dr Uri Hanani  
Department of Information and Library Studies, Bar-Ilan University  
Ramat-Gan  
Israel 52900  
Tel: +972 3 5318351  
Fax: +972 2 784368, +972 3 5347601  
E-mail: hanani@gilo.jlm.k12.il

## References

- [1] Schatz, B. and H. Chen (1996) Building large-scale digital library, *IEEE Computer*, **29**(5), May 96, 22–26.
- [2] Brewer, A. et al. (1996) The role of intermediary services in emerging digital libraries, *Proceedings of the 1st International Conference on Digital Libraries*, Bethesda, MD, March 1996, pp. 29–35.
- [3] Atkins, D.E. et al. (1996) Towards inquiry-based education through interacting software agents, *IEEE Computer*, **29**(5), 69–76.
- [4] Rao, R. et al. (1995) Rich interaction in the digital library, *Communications of the ACM*, **38**(4), 29–39.
- [5] Hanani, U. (1995) Intelligent hypertext: a frame for automatic search guidance system, *Proceedings of the 19th International Online Information Meeting*, London, pp. 219–229.
- [6] Bowman, C.M. et al. (1994) Scalable Internet resource discovery: research problems and approaches, *Communications of the ACM*, **37**(8), 98–107.
- [7] Bowman, C.M. et al. (1995) Harvest: a scalable, customisable discovery and access system, *Dept. Comput. Sci. Univ. Colorado, TR CU-CS-732-94*, March 1995.
- [8] Rada, R. (1991) *Hypertext: From Text to Expertext*, McGraw-Hill Book Company, London.
- [9] Chen, H. et al. (1993) Generating, integrating, and activating thesauri for concepts-based documents retrieval, *IEEE Expert*, 25–34.
- [10] Shapira, B., A. Raveh, P. Shoval and U. Hanani (1996) Hypertext browsing: a new model for information filtering based on user profile and data clustering, *Online & CDROM Review*, **20**(1), 3–10.
- [11] Gentner, D.R. and J. Grudin (1996) Design models for computer-human interface, *IEEE Computer*, **29**(6), 28–35.
- [12] Wiederhold, G. (1996) Digital libraries, value and productivity, *Communications of the ACM*, **38**(4), 85–96.

## Appendix

The structure of the DL is somehow a result of the physical layout of the Gilo school infrastructure. The computer network described as follows in Figure 3 and illustrated in Figure 4 may serve also to demonstrate to the reader the computing resources of this project.

### GILO HIGH SCHOOL COMPUTING RESOURCES

#### *Configuration Student Labs: Hardware*

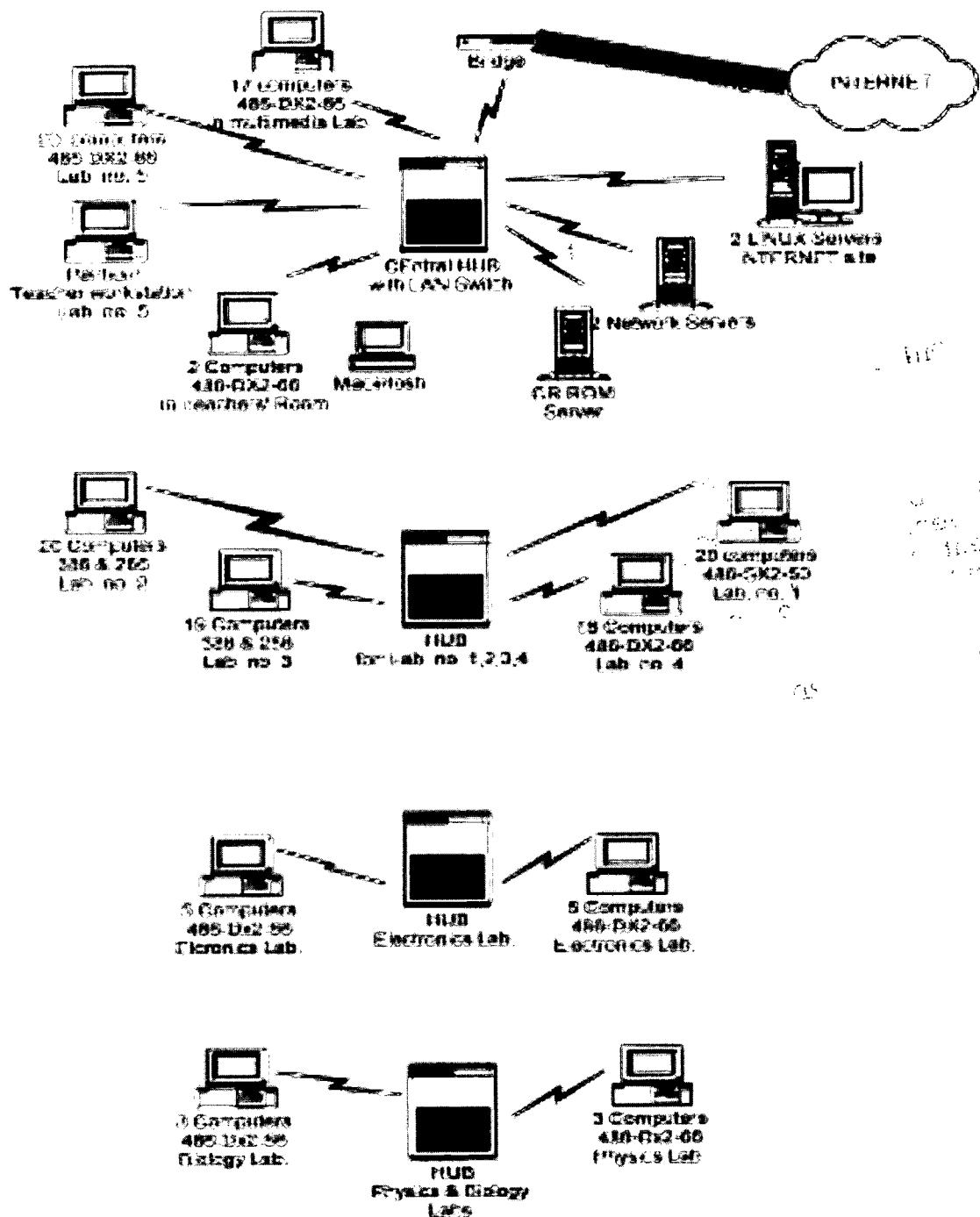
- 1) **Two Novell 3.12 File Servers:** (150+ workstations) serving 7 computer labs, 3 science labs and 10 administrative nodes. All are networked on an optical fiber backbone with a distributed star topology running through a 10BaseT cabling system. Novell 3.12 soon to be updated to version 4.1.
- 2) **Two secondary servers:** a) CD-ROM server working on a cd-rom tower of 8 drives, allowing concurrent sessions on single or multiple drives. b) EDUNETICS server of classroom learning sessions.
- 3) **Two Unix servers:** running on Intel 486DX and Pentium platforms (LINUX). They are connected to a dedicated 64k frame-relay line Internet connection containing: a) World Wide Web server that holds the Gilo High School web site [www.gilo.jlm.k12.il](http://www.gilo.jlm.k12.il) (bushwack) b) Mail Server (luke). c) "KATSIR" knowledge robot searching selected Internet sites on educational issues. d) four line dial-in modems - soon to be expanded to a multi-line modem Digiboard card.
- 4) **LANNET Intelligent Hub Switchers:** routing data packets on the ten segments of the network. The switcher is capable of: a) supporting combination of 10 and 100 Mb/sec network cards. Now we are running at 10 Mb/sec. b) supporting high bandwidth ATM technology for use in transferring voice and video packets. That is, the infrastructure ready to support video conferencing and multimedia presentations.
- 5) One **Print Server** spooling print jobs to 10 printers.

#### *Software configurations:*

- 1) Microsoft Windows and a full suite Microsoft application running from the file servers.
- 2) Microsoft, Novell, Linux, Borland application and software development products are running on individual stations with the aim of having them available network-wide in the future. Also Visual Basic, C++, and Turbo Pascal.
- 3) Multimedia and education authoring systems: ACTION, Astound, PowerPoint.
- 4) Two e-mail systems servers - one on the Unix server running SENDMAIL for staff and teacher use, the second on the Novell server running Mercury that connects students to internet e-mail.
- 5) World Wide Web authoring tools - HTML utilities that allows over 100 students to have home pages on the Internet.

**Figure 3:** Description of the Gilo High School computing resources.

## Schematic Layout of the Gilo Computers Network



**Remark:** All the HUBs are interconnected to one network with fiber optics.

**Figure 4:** Schematic layout of the Gilo High School computer network.



**U.S. DEPARTMENT OF EDUCATION**  
*Office of Educational Research and Improvement (OERI)*  
*Educational Resources Information Center (ERIC)*



## **NOTICE**

### **REPRODUCTION BASIS**



This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.



This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").